

Appln. No. 09/911,522
Amendment dated May 21, 2004
Reply to Office Action mailed November 21, 2003

REMARKS

Reconsideration is respectfully requested.

Claims 1-15 remain in this application. No claims have been cancelled. No claims have been withdrawn. Claims 16-20 have been added.

Preliminary Matters

The Incorporation By Reference listing beginning on page 2 of the specification as originally filed, has been amended to include the filing date and application serial number of each of the referenced co-pending patent applications.

In reviewing claim 1, the applicant has noticed an antecedent basis issue with the initial recital of the "candidate database listing". As such claim 1, has been amended to correct the antecedent basis issue in the recitation of the "candidate database listing". The applicant respectfully submits that the correction of this typographical error does not narrow the scope of the claim. The applicant also respectfully submits that the inserted line is fully supported by the specification and does not present new matter.

The Examiner's rejections will be considered in the order of their occurrence in the Office Action.

Paragraphs 1 and 2 of the Office Action

Claims 1-3, 5-6 and 8-15 have been rejected under 35 U.S.C. §102(b) as being anticipated by Wu (6,553,364).

The specification of the present application states in part that "[s]earch engines are known in the prior art and allow a user to search for sites that have some keyword corresponding to the user's query. While it is true that millions of documents are readily available as static pages to users through search engines, much more of the total content of the Internet, in the form of dynamic content, has remained relatively difficult to access

through more conventional search engine techniques." Page 3 line 31 through page 4 line 5. The Wu reference teaches one such search engine, namely Yahoo!, which provides search capabilities of static pages through the use of a hierarchical structure. This organized hierarchical structure then becomes a search database. Illustrative examples of this teaching of Wu include:

The present invention relates to an apparatus for searching for selected documents in a document repository containing a large number of documents. A search engine according to one embodiment of the present invention receives a search expression and, based on that search expression, searches for matching documents in the document repository and returns either the matching documents or a list of references to each of the matching documents. *Column 3 lines 33 through 39.*

The documents described in the main example herein are records in a search database. The search database is organized as a hierarchical structure of categories and site references. The structure might be automatically generated, but in the embodiment known as the Yahoo! search database, the categories and site references are placed in appropriate locations in the hierarchy by an editorial staff using the experience and suggestions from site submitters. *Column 3 lines 58 through 65.*

FIG. 2 shows how the documents in document repository 20 are logically arranged. In this example, documents are elements of a search database which is used to locate WWW sites of interest. Each document represents a topical category or a site and each document is shown as a record 38 in a hierarchical structure being in parent or child relation with other records. Each record 38 is shown with a document number 40 and content 46. In the case of a document which is a category, content 46 is the title of the category and other text (not shown), such as hidden keywords, synonyms, descriptions, etc., while the content of documents which refer to sites includes a title, a URL, a description, hidden keywords, synonyms, etc. Of course, some of these elements can be blank, where appropriate or desired. As explained above, in the Yahoo! search database, the documents are positioned in the hierarchical structure by an editorial staff. In a typical procedure, a site promoter will submit site information to the editorial staff, such as a site title, site URL, proposed location in the hierarchy, description, etc. The editorial staff then evaluates the submission,

changing the suggested location if a more appropriate location exists, cross links as needed, and adds, in some cases, hidden keywords, synonyms and/or a document importance weighting value. *Column 4 line 58 through column 5 line 14.*

Claim 1, not only as amended, but also as originally filed, includes the requirements of acquiring a listing of a plurality of qualified databases, submitting a query from the user to the plurality of databases, acquiring a collection of responsive content from the plurality of qualified databases, among others.

The Wu reference teaches a method of creating one such database to facilitate access of static information or documents to a user. The database of Wu might, for purposes of argument only, qualify as one of the databases to be queried by the present invention.

Therefore it is respectfully submitted that Wu fails to anticipate the present invention as recited in claim 1, and that claim 1, is allowable over Wu.

Claims 2-3, 5-6 and 8-15 are all dependent upon claim 1, which for the rationale discussed more fully above is believed to be allowable. Therefore, by virtue of their dependence from claim 1, claims 2-3, 5-6 and 8-15 are all believed to be allowable.

Regarding claim 2 specifically, the office action states in part that the Wu reference teaches "obtaining an inclusion list providing a plurality of terms and sources restricting associations for said step of acquiring responsive content (Wu col. 6 lines 6, lines 65-67)..."

The referenced sections of Wu are as follows:

"Document repository 20 includes the necessary processing logic to return documents requested by document number and either document repository 20 or search engine 36 contain processing logic to search a record for an instance of a field value which matches a query term." Column 6 lines 9 through 9.

“The use of document repository 20 and word index 22 will now be described with reference to an example. In this example, a user is searching for documents and presents a search request with a query string "The game of Go". Search engine 36 looks up each of the terms in word index 22. Because they are so common, "the" and "of" are either ignored by search engine 36 or word index 22 returns instructions to ignore those words, as described above. Search engine 38 then reads the document lists for "game" and "go", generates a match list for each term and applies an "AND" operator to the match lists as described below in connection with FIG. 6.” Column 6 line 61 through column 7 line 5.

However, Wu specifically teaches away from requiring the terms of an inclusion list to be found in the responsive content. Instead Wu teaches utilizing category labels to produce the results of a search. Wu continues the above referenced teaching as follows:

“A match list is a list of all the documents that contain the list's match tag either directly or indirectly. A match tag is a word or other search term or search element, depending on what the query term is. FIG. 2 illustrates why indirect matches are important. The example used throughout this description is a search for categories and sites related to the game of "Go", a well-known board game using black and white markers. Since the name of the game happens to be the same as a common word in the English language, searching for "go" would result in too many unrelated matches. **However, as can be seen, searching for "go" and "game" in the same document would result in no matches.**” Column 7 lines 6 through 17. Emphasis Added.

Therefore it is respectfully submitted that the WU reference not only fails to teach the use of an inclusion list as required by claim 2 as originally filed, but also teaches away from the present invention by teaching the use of category labels to augment search results.

With regard to claim 3, the WU reference teaches, at least in part, as follows:

FIG. 2 shows how the documents in document repository 20 are logically arranged. In this example, documents are elements of a search database which is used to locate WWW sites of interest. Each document represents a topical category or a site and each document is

shown as a record 38 in a hierarchical structure being in parent or child relation with other records. Each record 38 is shown with a document number 40 and content 46. In the case of a document which is a category, content 46 is the title of the category and other text (not shown), such as hidden keywords, synonyms, descriptions, etc., while the content of documents which refer to sites includes a title, a URL, a description, hidden keywords, synonyms, etc. Of course, some of these elements can be blank, where appropriate or desired. As explained above, in the Yahoo! search database, the documents are positioned in the hierarchical structure by an editorial staff. In a typical procedure, a site promoter will submit site information to the editorial staff, such as a site title, site URL, proposed location in the hierarchy, description, etc. The editorial staff then evaluates the submission, changing the suggested location if a more appropriate location exists, cross links as needed, and adds, in some cases, hidden keywords, synonyms and/or a document importance weighting value. Column 4 line 58 through column 5 line 15.

As can be observed from a reading of the plain text of WU, this section describes information related to a single document (which might be placed within a database) and not an initial page from the database or each one of the candidate databases as required by claim 3. Therefore it is respectfully submitted that Wu fails to teach suggest, or disclose the present invention, especially as disclosed in claim 3 as originally filed.

With regard to claim 8, the office action states that Wu Teaches "determining an input location for passing queries by said initial page to each one of said plurality of databases (Wu, Fig 5);..."

Figure 5 of Wu is a screen shot of a browser display of search results according to one embodiment. (Column 3 lines 21 through 23). More specifically, Figure 5 of Wu does not show, teach, suggest, or disclose a query input field, let alone the step of determining the location of such a field for submitting queries to a database.

Therefore it is respectfully submitted that Wu does not anticipate the present invention, especially as recited by claim 8 as originally filed.

With regard to claim 10 specifically, the office action states in part that Wu "further teaches removing duplicate pieces of responsive content (Wu, Col 6, lines 10-67)..." The cited portion of Wu is as follows:

Referring now to word index 22 shown in FIG. 3, a small subset of the contents of word index 22 is there shown. Word index 22 is organized as a plurality of records, with one record per word occurring in the documents of document repository 20, sorted in alphabetical order by word. Each record 52 in word index 22 is shown with a tag identifying the word, followed by a list of document numbers. These document numbers represent the list of documents containing the word. Where a word is so common as to be a search term of limited usefulness, such as the word "the", its record does not list all of the documents containing the word, but just an indication that the word should be ignored.

In the preferred embodiment, search engine 36 uses a document profile array 49 to improve search speed. Typically, array 49 is stored in memory for quick access. Array 49 has one record per document and each record includes fields for a document number, a document repository pointer, a time stamp, a child record range and an importance weighting value. The document numbers correspond to document numbers of document repository 20 and the document repository pointers correspond to physical disk locations of the documents in document repository 20, so that array 49 can be used to perform some operations on documents which don't require an access of document repository 20 itself. The time stamp identifies the time that the document was last modified. The range of children field indicates which records are below the instant document in the hierarchy, so that search engine 36 can quickly build a match list without having to refer to document repository 20 too often.

The importance weighting value is a value set automatically, or by an editorial staff, to indicate how valuable and/or relevant a particular category or site is relative to other categories and sites. The importance weighting value of a record might be adjusted based on external events or the significance of a site. For example, a site related to a particular group which is currently in the news might be given a higher weighting, or a site might be given a higher weighting if the editorial staff determines that the site is a popular or well-designed. Although the category records in array 49 shown in FIG. 3 do not have weighted categories, weighted categories might be useful.

For example, during boating season, the weighting for document #9 (category "boating"), might be increased. A record's weighting comes into play when multiple documents are being displayed as a search result, as the display documents are displayed in order by their weighting values. Of course, other weighting factors, as described below, might override the importance weighting or be combined with it to form an overall weighting.

The use of document repository 20 and word index 22 will now be described with reference to an example. In this example, a user is searching for documents and presents a search request with a query string "The game of Go". Search engine 36 looks up each of the terms in word index 22. Because they are so common, "the" and "of" are either ignored by search engine 36 or word index 22 returns instructions to ignore those words, as described above. Search engine 38 then reads the document lists for "game" and "go", generates a match list for each term and applies an "AND" operator to the match lists as described below in connection with FIG. 6.

As can be seen from a plain reading of the reference, Wu is silent as to duplicate pieces of responsive content, and does not teach, suggest, or disclose the removal of duplicate pieces of responsive content. Therefore it is again respectfully submitted that WU does not anticipate the present invention, especially as recited in claim 10 as originally filed.

With regard to claim 11, specifically, the office action states that Wu teaches "obtaining an inclusion list to restrict matches on predetermined words or sources (Wu column 6 lines 10 – column 11, line 47)...". The Wu reference in its entirety, and more specifically the 30 paragraphs cited in support of this assertion have been reviewed by the applicant and no suggestion of an inclusion list can be found. Further the plain reading of this section of Wu teaches away from the present invention by forcing the return of documents which fail to include all of the required terms through the use of artificially generated "category" tags to generate "indirect" matches.

Therefore it is respectfully submitted that the Wu reference fails to anticipate the present invention, especially as recited in claim 11, as originally filed.

With regard to claim 12, specifically, the present invention requires the step of saving the categorization file to a storage medium for use in searching the collection of responsive content. The office action cites the following portion of WU in support of the assertion that Wu teaches this requirement.

"In the case of a document which is a category, content 46 is the title of the category and other text (not shown), such as hidden keywords, synonyms, descriptions, etc., while the content of documents which refer to sites includes a title, a URL, a description, hidden keywords, synonyms, etc. Of course, some of these elements can be blank, where appropriate or desired. As explained above, in the Yahoo! search database, the documents are positioned in the hierarchical structure by an editorial staff. In a typical procedure, a site promoter will submit site information to the editorial staff, such as a site title, site URL, proposed location in the hierarchy, description, etc. The editorial staff then evaluates the submission, changing the suggested location if a more appropriate location exists, cross links as needed, and adds, in some cases, hidden keywords, synonyms and/or a document importance weighting value."

It is respectfully submitted that there is no suggestion or teaching to save the categorization file to a storage medium as purported by the office action. Therefore it is again respectfully submitted that the Wu reference fails to anticipate the present invention, especially as recited by claim 12 as originally filed.

With regard to claim 14, specifically, the present invention requires the step of developing an extract summary for each piece of responsive content. The office action cites the following portion of Wu in support of the assertion that WU teaches this requirement.

Referring now to word index 22 shown in FIG. 3, a small subset of the contents of word index 22 is there shown. Word index 22 is

organized as a plurality of records, with one record per word occurring in the documents of document repository 20, sorted in alphabetical order by word. Each record 52 in word index 22 is shown with a tag identifying the word, followed by a list of document numbers. These document numbers represent the list of documents containing the word. Where a word is so common as to be a search term of limited usefulness, such as the word "the", its record does not list all of the documents containing the word, but just an indication that the word should be ignored.

In the preferred embodiment, search engine 36 uses a document profile array 49 to improve search speed. Typically, array 49 is stored in memory for quick access. Array 49 has one record per document and each record includes fields for a document number, a document repository pointer, a time stamp, a child record range and an importance weighting value. The document numbers correspond to document numbers of document repository 20 and the document repository pointers correspond to physical disk locations of the documents in document repository 20, so that array 49 can be used to perform some operations on documents which don't require an access of document repository 20 itself. The time stamp identifies the time that the document was last modified. The range of children field indicates which records are below the instant document in the hierarchy, so that search engine 36 can quickly build a match list without having to refer to document repository 20 too often.

The importance weighting value is a value set automatically, or by an editorial staff, to indicate how valuable and/or relevant a particular category or site is relative to other categories and sites. The importance weighting value of a record might be adjusted based on external events or the significance of a site. For example, a site related to a particular group which is currently in the news might be given a higher weighting, or a site might be given a higher weighting if the editorial staff determines that the site is a popular or well-designed. Although the category records in array 49 shown in FIG. 3 do not have weighted categories, weighted categories might be useful. For example, during boating season, the weighting for document #9 (category "boating"), might be increased. A record's weighting comes into play when multiple documents are being displayed as a search result, as the display documents are displayed in order by their weighting values. Of course, other weighting factors, as described below, might override the importance weighting or be combined with it to form an overall weighting.

The use of document repository 20 and word index 22 will now be described with reference to an example. In this example, a user is

searching for documents and presents a search request with a query string "The game of Go". Search engine 36 looks up each of the terms in word index 22. Because they are so common, "the" and "of" are either ignored by search engine 36 or word index 22 returns instructions to ignore those words, as described above. Search engine 38 then reads the document lists for "game" and "go", generates a match list for each term and applies an "AND" operator to the match lists as described below in connection with FIG. 6.

It is respectfully submitted that there is nothing in the above extract which teaches, suggests or discloses the step of forming an extract summary as required by claim 14 as originally filed.

With regard to claim 15, each of the preceding arguments as applied to the individual claims as listed above are also applicable to claim 15, for the same respective rationales. Therefore, it is further respectfully submitted that the WU reference fails to anticipate the present invention.

Withdrawal of the §102(b) rejection of claims 1-3, 5-6 and 8-15 is therefore respectfully requested.

Paragraphs 3 and 4 of the Office Action

Claims 4 and 7 have been rejected under 35 U.S.C. Section 103(a) as being unpatentable over Wu (6,553,364).

Claims 4 and 7 are both dependent upon claim 1, which for the rationale discussed more fully above is believed to be allowable. Therefore, by virtue of their dependence from claim 1, claims 4 and 7 are both believed to be allowable.

Further, both claims 4 and 7 require the limitation of removing each candidate database with a relevancy below a minimum threshold value from qualification.

The office action correctly points out that the Wu reference does not teach this limitation. However, the office action interprets the weighting values within Wu as leading one to an obvious extension of Wu. The

Appln. No. 09/911,522

Amendment dated May 21, 2004

Reply to Office Action mailed November 21, 2003

applicant respectfully submits that this assertion by the office action is inconsistent with the plain language of Wu. The Wu reference states as follows:

“The importance weighting value is a value set automatically, or by an editorial staff, to indicate how valuable and/or relevant a particular category or site is relative to other categories and sites. The importance weighting value of a record might be adjusted based on external events or the significance of a site. For example, a site related to a particular group which is currently in the news might be given a higher weighting, or a site might be given a higher weighting if the editorial staff determines that the site is a popular or well-designed. Although the category records in array 49 shown in FIG. 3 do not have weighted categories, weighted categories might be useful. For example, during boating season, the weighting for document #9 (category "boating"), might be increased. A record's weighting comes into play when multiple documents are being displayed as a search result, as the display documents are displayed in order by their weighting values. Of course, other weighting factors, as described below, might override the importance weighting or be combined with it to form an overall weighting.”

As can be readily appreciated, nothing in Wu associated the weighting value with the removal of a document (or database), but rather as an editorial tool to change an order of presentation based on a range of factors.

Therefore it is respectfully submitted that the Wu reference fails to support a finding of obviousness under 35 USC 103(a) and that claims 4 and 7, as originally filed are allowable over Wu.

Withdrawal of the §103(a) rejection of claims 4 and 7 is therefore respectfully requested.

Finally, new claims 16 through 20 have been added to vary the scope of the present invention. All limitations within the claims are supported by the specification as originally filed, and therefore it is respectfully submitted that no new matter has been added.

Appln. No. 09/911,522

Amendment dated May 21, 2004

Reply to Office Action mailed November 21, 2003

CONCLUSION

In light of the foregoing amendments and remarks, early reconsideration and allowance of this application are most courteously solicited.

Respectfully submitted,

LEONARD & PROEHL, Prof. L.L.C.

By 

Date: 5/21/04

Mark A. Ekse (Reg. No. 50,130)

LEONARD & PROEHL, Prof. L.L.C.

Customer Number 40158

3500 South First Avenue Circle, Suite 250

Sioux Falls, SD 57105-5807

(605)339-2028 FAX (605)336-1931